Probability Box Propagation: Benchmarking Challenge Problems

J. Sadeghi, M. De Angelis, E. Patelli
Institute for Risk and Uncertainty, University of Liverpool, United Kingdom

T. Fetz, M. Oberguggenberger
Unit of Engineering Mathematics, University of Innsbruck, Austria

ABSTRACT

Probability boxes are an effective tool for the quantification of epistemic uncertainty in structural reliability analysis. However, it has been acknowledged that their propagation through a typical structural reliability analysis problem is in general non-trivial, due to the large number of samples of the limit state function required. The challenges are threefold: firstly minimising the computational time required, secondly robustness of the obtained bounds on the failure probability of the structure, and finally the tightness of the obtained bounds. Many methods have emerged in the literature which claim to solve the problem of efficient sampling, but in general the relative efficiency of the proposed methods is currently not well understood. It is likely that the efficiency and robustness of the methods is a function of the type of problem being studied. Therefore, we propose a range of test cases which can be applied to new algorithms which aim to solve this problem. The test cases consist of varying numbers of probability box inputs, rare and common failure events and linear and highly non-linear limit state functions.

1 INTRODUCTION

In recent years probability boxes have emerged as an effective tool for the quantification of epistemic uncertainty in structural reliability analysis (Patelli et al., 2015). However, it has been acknowledged that their propagation through a typical structural reliability analysis problem is in general non-trivial, due to the large number of samples of the limit state function required, as effectively multiple probability distributions must be sampled (De Angelis, 2015). Consider the following forward propagation problem:

$$\bar{P}_f = \max_{\theta} \int_{g(x)<0} p(x,\theta)dx, \tag{1}$$

where $\bar{P}_f$ is the upper bound on the failure probability, $g(x)$ is the well known limit state function which in general can be computationally expensive, and $p(x,\theta)$ is a parameterised probability density function where parameter θ falls within a specified interval. The challenges are threefold: firstly minimising the computational time required, secondly robustness of the obtained bounds on the failure probability of the structure, and finally the tightness of the obtained bounds.

Therefore, many methods have emerged in the literature which claim to solve the problem of efficient sampling. An early review of these methods is available in Bruns and Paredis (2006), however in the past ten years the state of the art has advanced considerably. Jiang et al. (2017) also review available methods, but do not focus on Monte Carlo Simulation and do not attempt to compare the relative efficiencies of the methods. The purpose of this conference paper is to invite researchers working in this area to contribute their algorithms to the challenge. Currently researchers from ETH Zurich, the University of Liverpool and the University of Innsbruck have agreed to contribute their algorithms to the challenge based on Multilevel Kriging (Schöbi and Sudret, 2017), Line Sampling (De Angelis, 2015), Scenario Optimisation (Sadeghi et al., 2018) and Importance Sampling (Fetz and Oberguggenberger, 2016; Fetz, 2017; Troffaes, 2017) respectively. It is known that double loop sampling methods are more effective in the case of distributional probability boxes and random set approaches are more effective in the case of distribution free probability boxes (Alvarez et al., 2017). However, in general, the relative efficiency of the proposed methods is currently not well understood. It is likely that the efficiency and robustness of the methods is a function of the type of problem being studied. Related test cases were proposed in Crespo et al. (2014), Oberkampf et al. (2004) and Ferson et al. (2004), however these test cases were designed to assess the ability of researchers to quantify their epistemic uncertainty and not to test the computational efficiency of the uncertainty propagation algorithms themselves. In fact, Ferson et al. (2004) acknowledged that a more computationally demanding set of test problems could prove useful.

Therefore, in Section 2 we propose a range of test cases which can be applied to new algorithms which aim to solve this problem. The test cases consist of varying numbers of probability box inputs, rare and common failure events and linear and highly non-linear limit state functions. The results obtained will be published in a journal paper in collaboration with contributing researchers which will allow software engineers and researchers to implement the appropriate method for their studies with confidence. After this we will make the data for the test cases freely available online so researchers in the future can apply the test cases to their algorithms in order for the efficiency of their algorithms to be immediately compared.

2 Test Cases

2.1 General comments

In order to allow ease of analysis of the test cases we will distribute Matlab (and Python, or C where requested) files containing the relevant functions. It is in our interests to make the test cases as accessible as possible to researchers, and therefore we have chosen to rely upon complex analytical functions instead of models using third party solvers, which may require licenses or take inordinate amounts of time to run. Instead, we simply request that researchers record the number of function calls which their software required. We also request that researchers specify if their function calls could have been reasonably parallelised, and how much time could feasibly be saved. For example, a brute force algorithm is clearly highly parallelisable despite being unreasonably time consuming. On the other hand an optimisation based algorithm could be very efficient, but might not result in a close to linear speedup when parallelised, as there is a dependence between samples.

2.2 Test Cases

We propose three test functions and inputs, where the inputs are contained in a vector x with components x_i where $i = 1,...,10$. The corresponding random variables x_i, $i = 1,\dots,10$ are independent and normally distributed with mean and variance in given intervals which describes a set of distributions. Alternatively, the p-box induced by this set of distributions can be considered (distribution free form). We would prefer that the problems are solved for both of these two models of uncertainty. The performance functions should be treated as black-boxes. For each problem we also include a simplified version with two inputs, which will serve the purpose of allowing the main features of the performance function to be visualised in two dimensions. Specifically the distributions which contribute to the maximum and minimum probability of failure are shown.

For each problem the distributions used are given in Table 1. For the simple performance function only x_1 and x_2 are considered.

We also include an Oscillator performance function, which is an existing test case used to benchmark algorithms for structural reliability analysis, as a fourth test problem.

Table 1: Input Variables for Problem 1

Parameter	Distribution	Mean Value	Standard deviation
x_1	Normal	[-1,1]	[1.5,2]
x_2	Normal	[-1,1]	[1.5,2]
x_3	Normal	[-1,1]	[1.5,2]
x_4	Normal	[-1,1]	[1.5,2]
x_5	Normal	[-1,1]	[1.5,2]
x_6	Normal	[-1,1]	[1.5,2]
x_7	Normal	[-1,1]	[1.5,2]
x_8	Normal	[-1,1]	[1.5,2]
x_9	Normal	[-1,1]	[1.5,2]
x_{10}	Normal	[-1,1]	[1.5,2]

2.2.1 Problem 1

The problem addresses a piecewise linear limit state surface, which in the simplified case is obtained from the performance function

$$g(x) = 10 - \max((1 + (x_1 + 5) + 2x_2), (3 + 2x_1 + x_2)). \tag{2}$$

Figure 1 shows the distributions which contribute to the maximum ($\mu_1 = \mu_2 = 1$, $\sigma_1 = \sigma_2 = 2$) and minimum ($\mu_1 = \mu_2 = -1$, $\sigma_1 = \sigma_2 = 1.5$) failure probability of Eqn. 2, which were obtained analytically.

The high dimensional performance function is given by

$$g(x) = 25 - \max\left((1 + x_1 + 5 + 2x_2 - x_3 + 3x_5 + x_6 + 0.2x_9 + 3x_{10}), (3 + 2x_1 + x_2 + 0.5x_4 + x_7 + 2x_8)\right), \tag{3}$$

which yields a low value of P_f.

2.2.2 Problem 2

The problem addresses the case of a nonlinear (quadratic), but smooth limit state surface, which in the simple case is obtained from the performance function

$$g(x) = 50 - \begin{pmatrix} x_1 - 2 & x_2 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} x_1 - 2 \\ x_2 \end{pmatrix}. \tag{4}$$

Figure 2 shows the distributions which contribute to the maximum ($\mu_1 = \mu_2 = -1$, $\sigma_1 = \sigma_2 = 2$) and minimum ($(\mu_1, \mu_2) \approx (1, 1)$, $\sigma_1 = \sigma_2 = 1.5$) failure probability of Eqn. 4, which were obtained by double loop Monte Carlo simulation with brute force grid search for the outer loop. The analytical solution for the mean corresponding to the lower probability is the rightmost intersection point of the square and axis of the ellipsoid. This example is somewhat similar to the example of a beam supported on both ends and bedded on two springs with uncertain spring constants given in Fetz (2017).

The high dimensional performance function is given by

$$g(x) = c - (x - v)^{\mathrm{T}} A (x - v), \tag{5}$$

where $v = -\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$ is a vector, $c = 350$ is a constant, and A is a matrix given by

$$A = \begin{pmatrix} 1.0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 \\ 0.9 & 1.0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 \\ 0.8 & 0.9 & 1.0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 \\ 0.7 & 0.8 & 0.9 & 1.0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 \\ 0.6 & 0.7 & 0.8 & 0.9 & 1.0 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 \\ 0.5 & 0.6 & 0.7 & 0.8 & 0.9 & 1.0 & 0.9 & 0.8 & 0.7 & 0.6 \\ 0.4 & 0.5 & 0.6 & 0.7 & 0.8 & 0.9 & 1.0 & 0.9 & 0.8 & 0.7 \\ 0.3 & 0.4 & 0.5 & 0.6 & 0.7 & 0.8 & 0.9 & 1.0 & 0.9 & 0.8 \\ 0.2 & 0.3 & 0.4 & 0.5 & 0.6 & 0.7 & 0.8 & 0.9 & 1.0 & 0.9 \\ 0.1 & 0.2 & 0.3 & 0.4 & 0.5 & 0.6 & 0.7 & 0.8 & 0.9 & 1.0 \end{pmatrix}.$$

2.2.3 *Problem 3*

The problem has a highly non-linear limit state surface which contains discontinuities; in the simple case it is obtained from the performance function

$$g(x) = \begin{cases} \frac{\tan(x_1)}{(x_2+1)(x_1+2)} + \frac{1}{4} & \text{if } x_2 < 5 \\ \frac{1}{4} & \text{otherwise} \end{cases} \tag{6}$$

which is displayed in Figure 3, where the distributions which contribute to the maximum and minimum failure probability were obtained by double loop Monte Carlo simulation with brute force grid search for the outer loop.

The high dimensional performance function is given by

$$g(x) = \begin{cases} (x_1 - \pi/2)\tan(x_1 - \pi/2) + 5\sin(x_2) + 10\sqrt{x_3^2 + x_4^2 + x_5^2} - 80 \\ \text{if } (x_{10} > 3 \vee x_{10} < -3) \wedge (x_9 > 0) \wedge (x_6 > 0) \\ 1 \text{ otherwise} \end{cases}$$

which yields a low P_f.

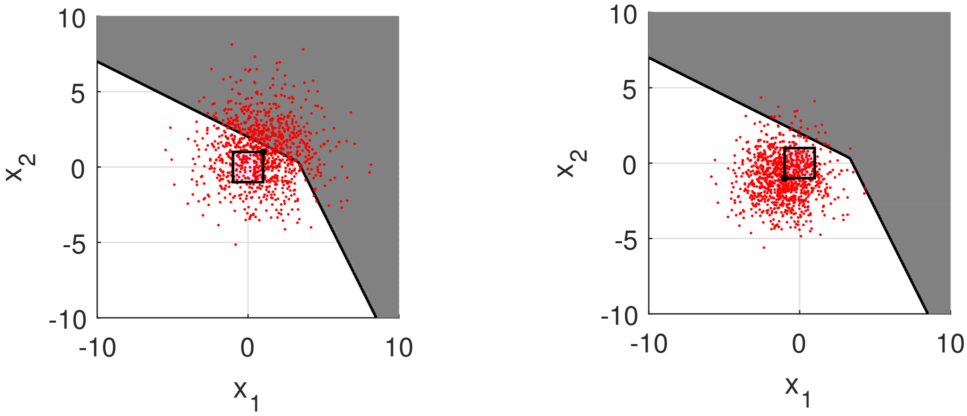
2.2.4 *Problem 4: Oscillator*

In the paper by Echard et al. (2013), two benchmarks are proposed to compare different algorithms for the computation of the probability of failure. The one of choice in this paper describes the performance of an undamped oscillator due to an impulse of length T_1.

The performance function is given by the following equation:

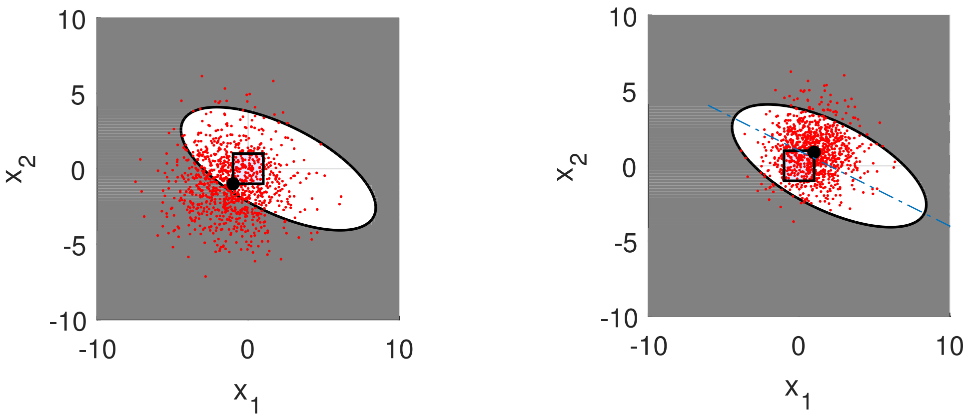
$$P_2 = 3R - \left| \frac{2F_1}{M\omega_0^2} \sin\left(\frac{\omega_0 T_1}{2}\right) \right|, \qquad \omega_0 = \sqrt{\frac{C_1 + C_2}{M}} \tag{7}$$

where the description and distributions of the six independent input random variables $X = \{C_1, C_2, M, R, F_1, T_1\}$ are given in Table 2.



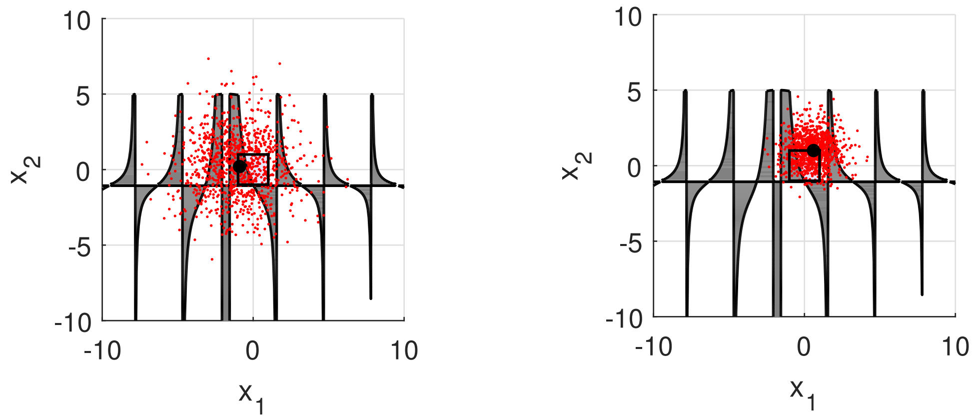
(a) Distribution corresponding to $\bar{P}_f = 0.439$ **(b)** Distribution corresponding to $\underline{P_f} = 0.026$

Figure 1: Illustration of extreme distributions for piecewise linear limit state function. The blue box represents interval imprecision on the mean parameter of the random variables and the bold point the mean corresponding to the upper/lower probability of failure. The red points are samples from the random variables. The shaded region represents the failure domain.



(a) Distribution corresponding to $\bar{P}_f = 0.433$ **(b)** Distribution corresponding to $\underline{P_f} = 0.071$

Figure 2: Illustration of extreme distributions for quadratic limit state function. The box represents interval imprecision on the mean parameter of the random variables and the bold point the mean corresponding to the upper/lower probability. The points are samples from the random variables. The shaded region represents the failure domain. The dashed blue line represents the semi-major axis of the ellipsoid.



(a) Distribution corresponding to $\bar{P}_f = 0.36$ **(b)** Distribution corresponding to $\underline{P_f} = 0.187$

Figure 3: Illustration of extreme distributions for discontinuous limit state function. The box represents interval imprecision on the mean parameter of the random variables and the bold point the mean corresponding to the upper/lower probability. The points are samples from the random variables.

Table 2: Parameters of the Gayton oscillator model in Echard et al. (2013)

Name	Description	Distribution	Statistics
C_1	First spring constant	Gaussian	$\mu_{C_1} = 1, COV_{C_1} = 10\%$
C_2	Second spring constant	Gaussian	$\mu_{C_2} = 0.1, COV_{C_2} = 10\%$
M	Oscillator mass	Gaussian	$\mu_M = 1, COV_M = 5\%$
R	Yield displacement	Gaussian	$\mu_R = [0.490.51], \sigma = 0.05$
F_1	Force acting on oscillator	Gaussian	$\mu_{F_1} = [-0.20.2], \sigma = 0.5$
T_1	Time interval the force is acting on the oscillator	Gaussian	$\mu_{T_1} = [0.95, 1.05], \sigma = 0.2$

3 Conclusion

We have proposed a set of functions we feel could emulate likely models in industrial use cases. We look forward to the response of researchers to these challenging benchmarking problems, and encourage interested parties to contact us for more information, if required.

Acknowledgements

We gratefully acknowledge the contributions of our partners and the University of Innsbruck and ETH Zurich.

References

Alvarez, D. A., J. E. Hurtado, and J. Ramírez (2017). Tighter bounds on the probability of failure than those provided by random set theory. *Computers & Structures 189*, 101–113.

Bruns, M. and C. J. Paredis (2006). Numerical methods for propagating imprecise uncertainty. In *Proceedings of the ASME International Design Engineering Conferences and Computers in Information Engineering Conference, Philadelphia, PA, Paper No. DETC2006-99237*.

Crespo, L. G., S. P. Kenny, and D. P. Giesy (2014). The nasa langley multidisciplinary uncertainty quantification challenge. In *16th AIAA Non-Deterministic Approaches Conference*, pp. 2014–1347. AIAA Reston, VA.

De Angelis, M. (2015). *Efficient Random Set Uncertainty Quantification by means of Advanced Sampling Techniques*. Ph. D. thesis, University of Liverpool.

Echard, B., N. Gayton, M. Lemaire, and N. Relun (2013). A combined importance sampling and Kriging reliability method for small failure probabilities with time-demanding numerical models. *Reliab. Eng. Syst. Safety 111*, 232–240.

Ferson, S., C. A. Joslyn, J. C. Helton, W. L. Oberkampf, and K. Sentz (2004). Summary from the epistemic uncertainty workshop: consensus amid diversity. *Reliability Engineering & System Safety 85*(1), 355–369.

Fetz, T. (2017). Efficient computation of upper probabilities of failure. In D. M. F. Christian Bucher, Bruce R. Ellingwood (Ed.), *Safety, Reliability, Risk, Resilience and Sustainability of Structures and Infrastructure*, pp. 493–502. 12th International Conference on Structural Safety & Reliability: TU-Verlag Vienna.

Fetz, T. and M. Oberguggenberger (2016). Imprecise random variables, random sets, and monte carlo simulation. *International Journal of Approximate Reasoning 78*, 252–264.

Jiang, C., J. Zheng, and X. Han (2017, Nov). Probability-interval hybrid uncertainty analysis for structures with both aleatory and epistemic uncertainties: a review. *Structural and Multidisciplinary Optimization*.

Oberkampf, W. L., J. C. Helton, C. A. Joslyn, S. F. Wojtkiewicz, and S. Ferson (2004). Challenge problems: uncertainty in system response given uncertain parameters. *Reliability Engineering & System Safety 85*(1), 11–19.

Patelli, E., D. A. Alvarez, M. Broggi, and M. de Angelis (2015). Uncertainty management in multidisciplinary design of critical safety systems. *Journal of Aerospace Information Systems 12*, 140–169.

Sadeghi, J., E. Patelli, and M. de Angelis (2018). Robust propagation of probability boxes by interval predictor models. Florianopolis, SC, Brazil. Proceedings of the joint ICVRAM ISUMA UNCERTAINTIES conference.

Schöbi, R. and B. Sudret (2017). Structural reliability analysis for p-boxes using multi-level metamodels. *Probabilistic Engineering Mechanics 48*, 27–38.

Troffaes, M. C. M. (2017, 10–14 Jul). A note on imprecise monte carlo over credal sets via importance sampling. In A. Antonucci, G. Corani, I. Couso, and S. Destercke (Eds.), *Proceedings of the Tenth International Symposium on Imprecise Probability: Theories and Applications*, Volume 62 of *Proceedings of Machine Learning Research*, pp. 325–332. PMLR.